

# MANOVA: Type I Error Rate Analysis

A Senior Project

Presented to

The Faculty of the Statistics Department

California Polytechnic State University, San Luis Obispo

In Partial Fulfillment

Of the Requirements for the Degree

Bachelor of Science

By

Kyle Wesley Gasperik

June, 2010

© 2010 Kyle Wesley Gasperik

## Contents

Introduction.....	3
What it is MANOVA & Type I Error .....	4
MANOVA.....	4
Type I Error.....	5
Assumptions.....	5
Independence .....	5
Distribution .....	5
Homogeneity of Variance .....	5
Simulations .....	6
Violating Assumptions.....	6
Analysis of Empirical Type I Error Rate .....	7
Multivariate Normal.....	7
Simulation I.....	7
Simulation II .....	9
Simulation III.....	10
Simulation IV.....	11
Multivariate Uniform .....	12
Simulation V .....	12
Simulation VI.....	14
Functions for Multivariate Methods .....	16
Future Projects .....	17
Conclusion .....	18
Appendix.....	19
Section A - Simulation Function.....	19
Section B - Other Simulations .....	20
Section C - Hotellings $T^2$ Functions .....	21
Section D - K Sample Profile Analysis Function .....	22

## Introduction

Multivariate analysis of variance (MANOVA) is an analysis method used mainly in the field of biostatistics, along with other scientific related fields. My project was introduced to me in order to help in-experienced programmers with the R software needed to perform this analysis. The multivariate statistical methods class at Cal Poly (Stat 419/ BIO 542) uses the R software but the class on R is not a pre-requisite. This makes it difficult for some students to use the software, and why I wrote some functions that would make analysis easier.

As my project progressed, questions started rising about some the assumptions of MANOVA and how certain settings would change variability in the results. The bulk of my senior project was performing simulations to try and understand how robust MANOVA is against violations of assumptions. Using the R software I looked at how introducing correlation between dependent variables, changing the covariance structure, would affect the empirical type I error rate. I also investigated if simulating multivariate data from a uniform distribution rather than a normal distribution affected this error rate. Type I error rate is very easy to calculate and interpret; this is why I used it to assess my simulations.

I wrote ten functions to perform a total of seventy individual simulations. Each simulation was then saved and results were graphed for analysis purposes. The body of my report highlights the simulations that were of most interest. Other simulations not included in the body can be found in the appendix.

## What it is MANOVA & Type I Error

### MANOVA

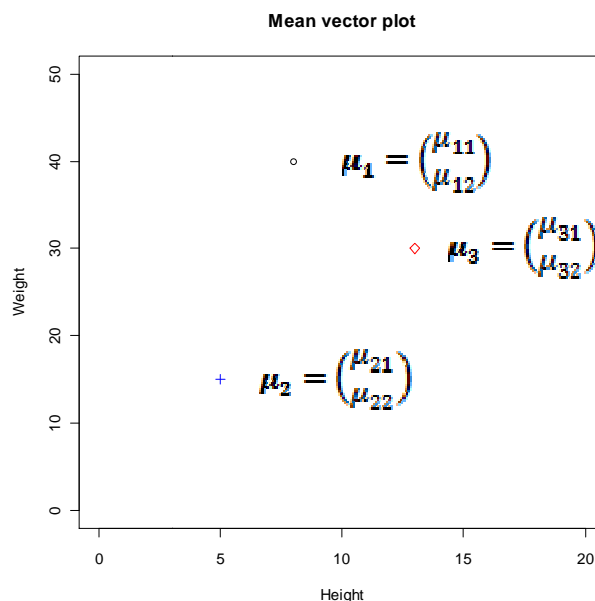
Multivariate Analysis of Variance (MANOVA) is a generalized extension of univariate analysis of variance (ANOVA). This analysis is used in situations where there are two or more dependent variables. MANOVA can also be thought of as a vector extension of ANOVA, so instead of testing if three or more means are equal for one dependent variable (ANOVA), MANOVA tests if three or more mean vectors are identical. For example, if someone wanted to test if the mean height and mean weight of three different football team's players are identical, one would use MANOVA. The null and alternative hypotheses are below in Table 1.

**Table 1: MANOVA Notation**

Term	Definition
$\mu_i$	mean weight and length vector for team i
$\mu_{i1}$	mean weight for team i
$\mu_{i2}$	mean length for team i
<b>Null Hypothesis</b>	$\mu_1 = \mu_2 = \mu_3$ or $\begin{pmatrix} \mu_{11} \\ \mu_{12} \end{pmatrix} = \begin{pmatrix} \mu_{21} \\ \mu_{22} \end{pmatrix} = \begin{pmatrix} \mu_{31} \\ \mu_{32} \end{pmatrix}$
<b>Alternative Hypothesis</b>	At least two $\mu_i$ not equal

In order to further understand what MANOVA is attempting to do, we can look at Figure 1, which graphically explains MANOVA. In the graph you will see three points plotted. These three points are to represent the three population mean vectors of weight and height for each team. MANOVA is trying to figure out if these points are far enough apart to say that two vectors jointly differ from one another.

**Figure 1: Graphical Representation of MANOVA**



## Type I Error

A type I error occurs when the null hypothesis is rejected when it is actually true. We are making an error because the null hypothesis is true but we say that it is false. The probability of committing a type I error is sometimes referred to as the significance level, denoted alpha. We set this value called alpha to 0.05 in most statistical test. Meaning, we only want to make this error five percent of the time. For all my comparisons I will be using this nominal alpha level of 0.05.

## Assumptions

With any statistical test there are assumptions that need to be satisfied in order for the test to be valid. With MANOVA, there are three main assumptions that need to be validated, which are listed below.

### Independence

The first assumption that's needs to be satisfied for MANOVA to be valid is that the observations have to be independent of each other. This clarifies that there is no pattern in the data and the sample is completely random.

### Distribution

We clarify that the multivariate samples are drwan from a Multivariate Normal Distribution.

### Homogeneity of Variance

The final assumption is that each group has equal population covariance structures. Now that we have multiple dependent variables we not only have to make sure that their variances are equal across all groups, but we also have to check to see if all the covariance's between them are equal. This is done using a covaraiance matrix. The diagonals of the matrix are the variance of each dependent variable, and the off-diagonals correspond to the covariance between different dependent variables. Figure 2. is a graphical explanation of the covariance matrix.

Figure 2: Covariance Matrix

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_2^2 & & & \vdots \\ \sigma_{31} & & \sigma_3^2 & & \vdots \\ \vdots & & & \ddots & \vdots \\ \sigma_{p1} & \cdots & \cdots & \cdots & \sigma_p^2 \end{bmatrix}$$

On the main diagonal of the matrix we have the variance of all of our variables. The off-diagonal contains the covariance between each variable. This covariance structure will be of dimension P x P, where P is the number of dependent variables you are testing.

## Simulations

A big part of my investigation had to do with using some statistical software. Below are some sections that help explain exactly how I performed my simulations using R.

### Violating Assumptions

The main goal of my investigation was to see how robust MANOVA was against certain violations of assumptions. I focused on two main violations. I wanted to see how robust MANOVA was when the dependent variables had different correlations between them in different groups and when we sampled from a Multivariate Uniform distribution. In order to do this analysis many functions were created, but first I had to decide which settings I wanted to narrow my test by.

There are many different simulation settings that can be changed. The main three are: the number of groups, the number of dependent variables, and the sample size within each group. For my investigation I can also change correlations between dependent variables for each group. As you will see in the next section, I limited which sets of correlations I looked at to try to find the best results.

For all of my simulations I stuck to three groups, and I looked at results for two, three, and four dependent variables. I then varied sample size in each group, the correlation between variables in each group, and the distribution that I sampled from (Normal or Uniform). When sampling, I made sure that each group had the exact same means, so I could then make an accurate decision when an error occurred.

All of my simulations investigated Type I error rate for MANOVA. I choose to use this measure for testing robustness because it was very easy to simulate and understand. An outline of how my simulation functions worked is found in Figure 3 below.

**Figure 3: Simulation Scheme**

1. Simulated specific settings
2. Randomized and ran a MANOVA
3. Compared F-Statistic from simulation against a 95 <sup>th</sup> percentile F-Statistic with the same degrees of freedom
4. Repeated this 1000 times
5. Calculated the proportion of times the random F-Stat was greater than the corresponding 95 <sup>th</sup> percentile.

A copy of one of my simulation functions can be found in section A of the appendix.

## Analysis of Empirical Type I Error Rate

In the following section I will present to you the different simulations that I performed to investigate type I error rate for MANOVA. Each simulation will first be described by what assumptions were violated and then a graph of empirical type I error rates to analyze. There are few other simulations not in the body of my paper that can be referenced in section B of the appendix. Table 2 shows the settings that were simulated for each different simulation.

**Table 2: Simulation Overview**

Simulation	Sample Size per Group (n)	# of Variables per Group (p)	Correlation Settings	Multivariate Distribution
<b>I</b>	10,30,50,100	2,3,4	Identical between pairs of variables w/in each group. Identical across 3 groups, varies from 0.05-0.95 in steps of 0.05	Normal
<b>II</b>	Varies from 10-100 in steps of 10	2,3,4	Identical between pairs of variables w/in each group. Varies largely across 3 groups	Normal
<b>III</b>	Varies from 10-200 in steps of 5	2,3,4	Identical between pairs of variables w/in each group. Cor1=0.5, Cor2=0.5, Cor=0.95	Normal
<b>IV</b>	Varies from 10-200 in steps of 10	2,3,4	Identical between pairs of variables w/in each group. Varies minimally across 3 groups	Normal
<b>V</b>	Varies from 10-200 in steps of 10	2,3,4	Variables are not correlated	All groups Uniform (0,1)
<b>VI</b>	Varies from 10-400 in steps of 10	2,3,4	Variables are not correlated	Group1=Uniform(0,6) Group2=Uniform(1,5) Group3=Uniform(2,4)

## Multivariate Normal

In the first set of simulations I ran, data was simulated from a multivariate normal distribution. Using R, I simulated a specific number of observations for each group from a multivariate normal distribution in which all groups have the same population mean vector. From there I started manipulating the covariance structures for each group. Finally I plotted all of the empirical type I error rates in order to see if any of these settings increased the type I error of the test.

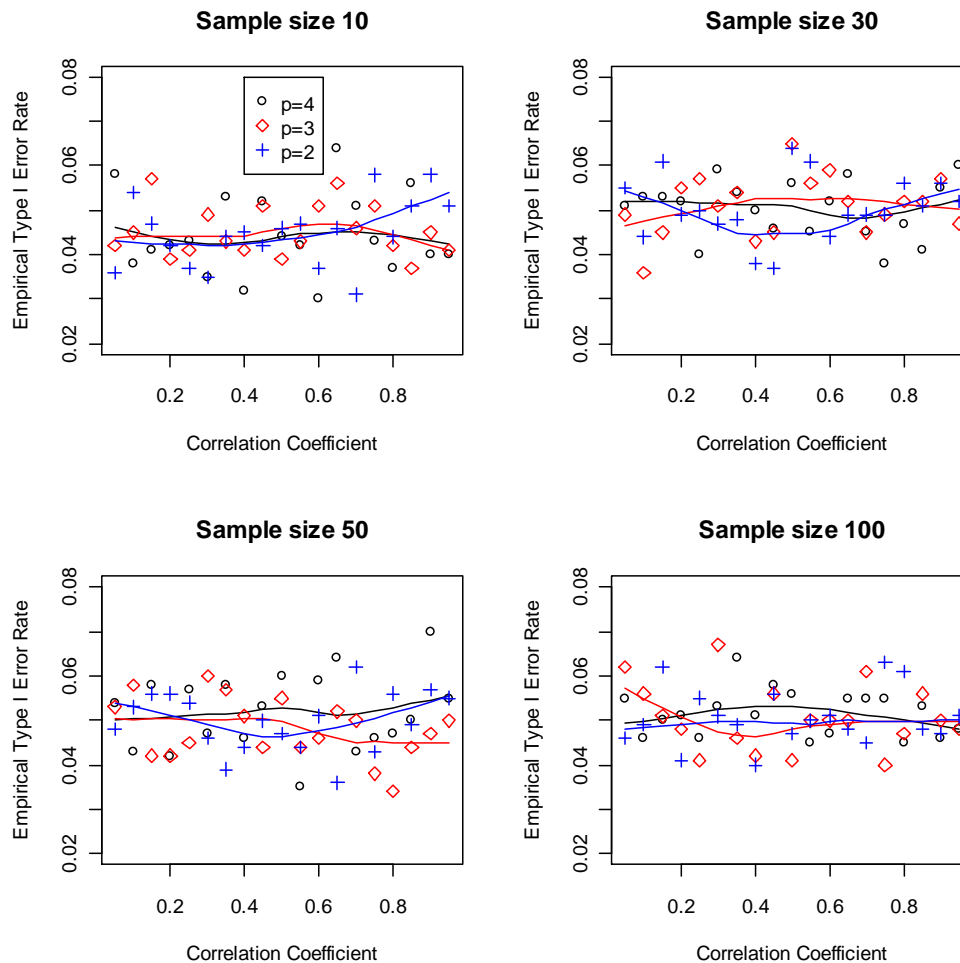
### Simulation I

This simulation investigated how the type I error rate was effected when dependent variables were correlated. Figure 4 contains four graphs of empirical type I error rate. On the x-axis is the value of the correlation between dependent variables for each group, and each graph corresponds to a different sample size, which is labeled on the top of the graph. As you can see

by the legend, the black, red, and blue points correspond to four, three and two dependent variables respectively.

In this simulation there are no technical violations of assumptions. Although the dependent variables are correlated, all of the covariance structures are equal because there is the same amount of correlation between variables for each group. Given that there is no violation of assumptions we would expect the average type I error rate to reach our nominal  $\alpha=0.05$  level. It turns out that all the average empirical type I error rates didn't stray from our nominal alpha level, and the graphs above seem to support to this fact.

**Figure 4: Empirical Type I Error Rate vs. Correlation**

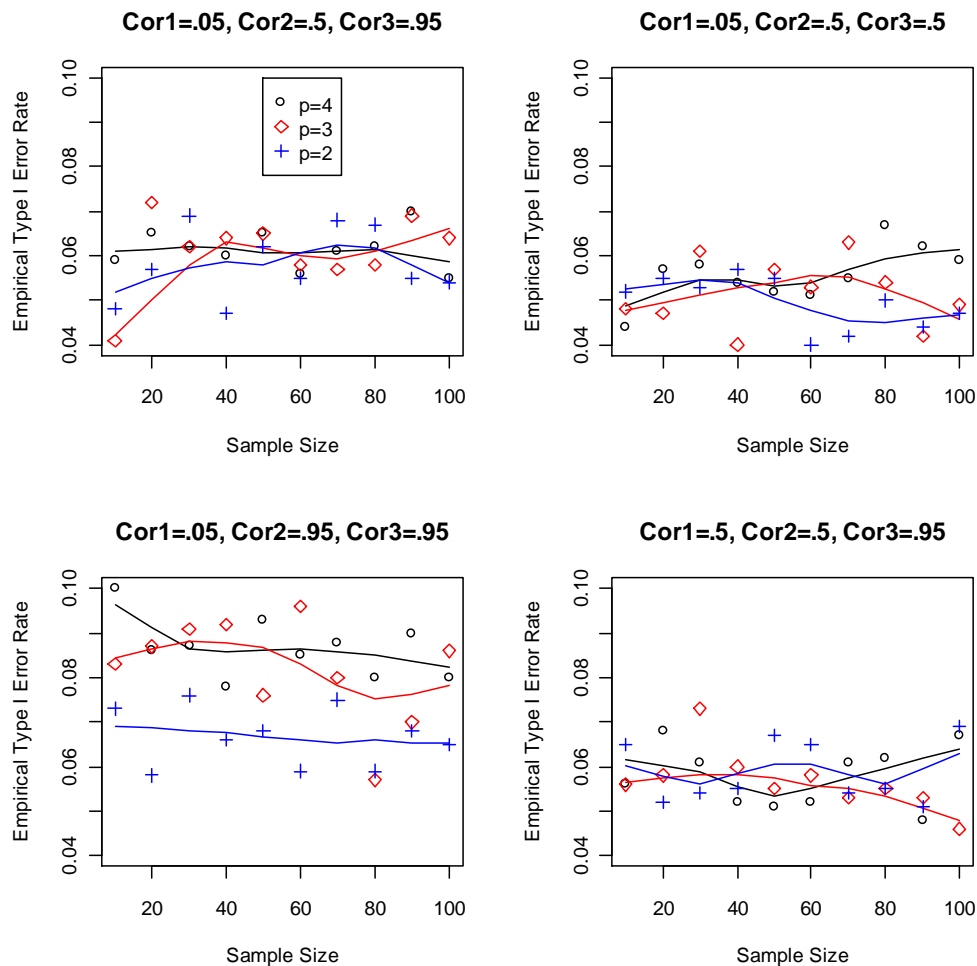




## Simulation II

Now that we have investigated groups with the same correlation between variables for each group, the next step was to change the correlations between variables for each group. Now we are violating the assumption of equal covariance structures, and we introduced correlation between dependent variables. Because of this violation we would expect our empirical type I error rates to be above the nominal level, which is shown by the graphs below. In this next simulation I investigate four different sets of correlation coefficients. In Figure 5 you will see four graphs; at the top of each graph are the correlation coefficients for each group labeled cor1, cor2, and cor3. On the x-axis is the different sample size in each group to investigate if increasing sample size might allow the type I error rate to reach our nominal level.

Figure 5: Changing Variable Correlation between Dependent Variable



All of the average empirical type I error rates are above the nominal level, but specifically, I will highlight the correlation set in which we saw the biggest difference. In the bottom left graph, the correlations are set to 0.05, 0.95, and 0.95. These settings produced average empirical type I error rates of 0.0867, 0.818, and 0.0667 for four, three, and two

dependent variables respectively. As you can see these error rates are well above our nominal level. Please refer to Table 3 for average empirical type I error rates.

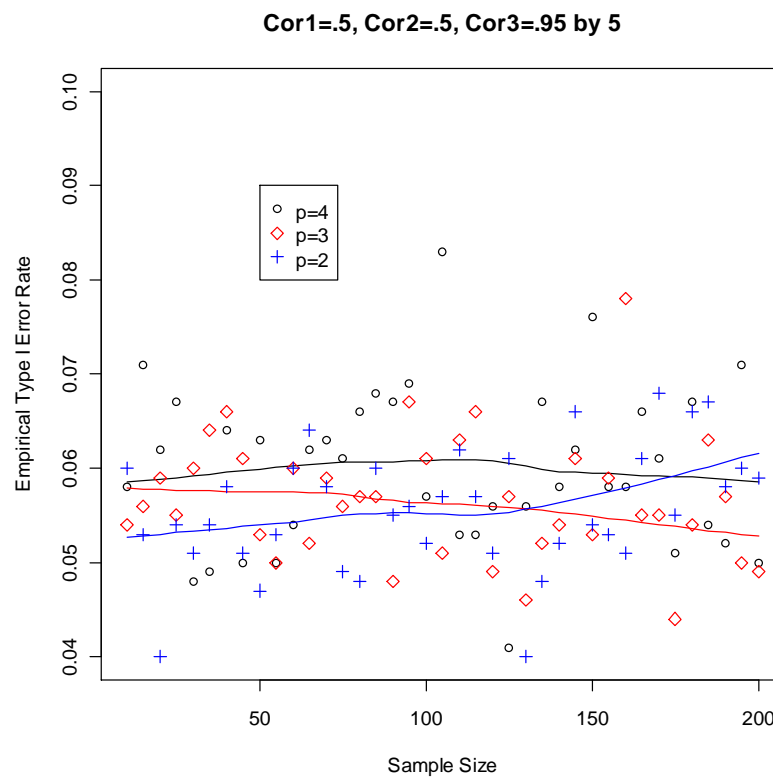
**Table 3: Simulation II Results**

Correlation Coefficients	Average Empirical Type I Error Rate 4 variables -- 3 variables -- 2 variables			
<b>Cor1=0.05,Cor2=0.5,Cor3=0.95</b>	0.0615	--	0.061	-- 0.0582
<b>Cor1=0.05,Cor2=0.5,Cor3=0.5</b>	0.0559	--	0.0514	-- 0.0495
<b>Cor1=0.05,Cor2=0.95,Cor3=0.95</b>	0.0867	--	0.0818	-- 0.0667
<b>Cor1=0.5,Cor2=0.5,Cor3=0.95</b>	0.0578	--	0.0567	-- 0.0587

### Simulation III

In order to investigate the effect of increasing sample size on empirical type I error rate, the next simulation replicated a correlation set but looked at more sample sizes. Figure 6 shows type I error rate vs. sample size when the correlations are set to 0.5, 0.5, 0.95, but now the sample size increases from 10-200 in steps of 5 rather than from 10-100 in steps of 10. This will allow us to investigate if an increasing sample size would eventually lead to a nominal alpha level for the MANOVA test even though we have violated some assumptions.

**Figure 6: Investigating Sample Size**



As seen in the graph above, we never quite reach our nominal alpha level. The average empirical type one error rates can be found in Table 4. Also, we are looking for any downward trend in the points, which there seems to be a slight one here, but nothing significant. However, there is still evidence that when dependent variables are correlated, the average type I error rate can be above our nominal level.

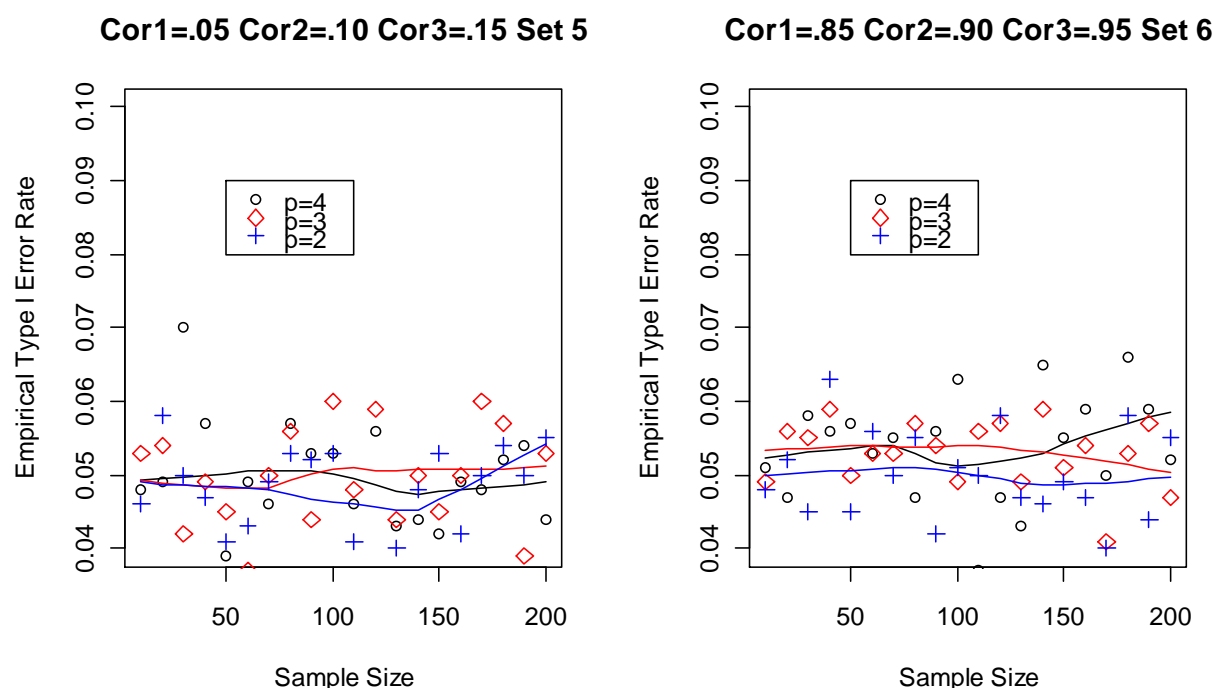
**Table 4: Simulation III Results**

Number of Variables	Average Empirical Type I Error Rate
4	0.060
3	0.0567
2	0.0556

### Simulation IV

Now that we have looked at groups with highly differing correlation coefficients, we thought it would be interesting to look at a simulation in which the correlations between variables for each group were very close together. In this simulation we looked at two other correlation sets. These correlations can be found above each graph with labeled Cor1, Cor2, and Cor3. Figure 7 shows two graphs of two different correlation sets that were investigated.

**Figure 7: Similar Correlation**



Looking at the graph on the left with the smaller correlation coefficients, it doesn't seem that the empirical type I error rates stray from 0.05. The average empirical type I error rates can be found in Table 5 on the following page. However, when the correlations are high between the

variables in each group, we did see some rates above our nominal level. All three error rates are barely above our nominal level. This is demonstrating when there is similar high correlation between dependent variables for each group the empirical type one error rate can be above our nominal level.

**Table 5: Simulation IV Results**

Correlation Coefficients	Average Empirical Type I Error Rate				
	4 variables -- 3 variables		3 variables -- 2 variables		2 variables
<b>Cor1=0.05,Cor2=0.10,Cor3=0.15</b>	0.0499	--	0.0495	--	0.0478
<b>Cor1=0.85,Cor2=0.90,Cor3=0.95</b>	0.0538	--	0.05295	--	0.0501

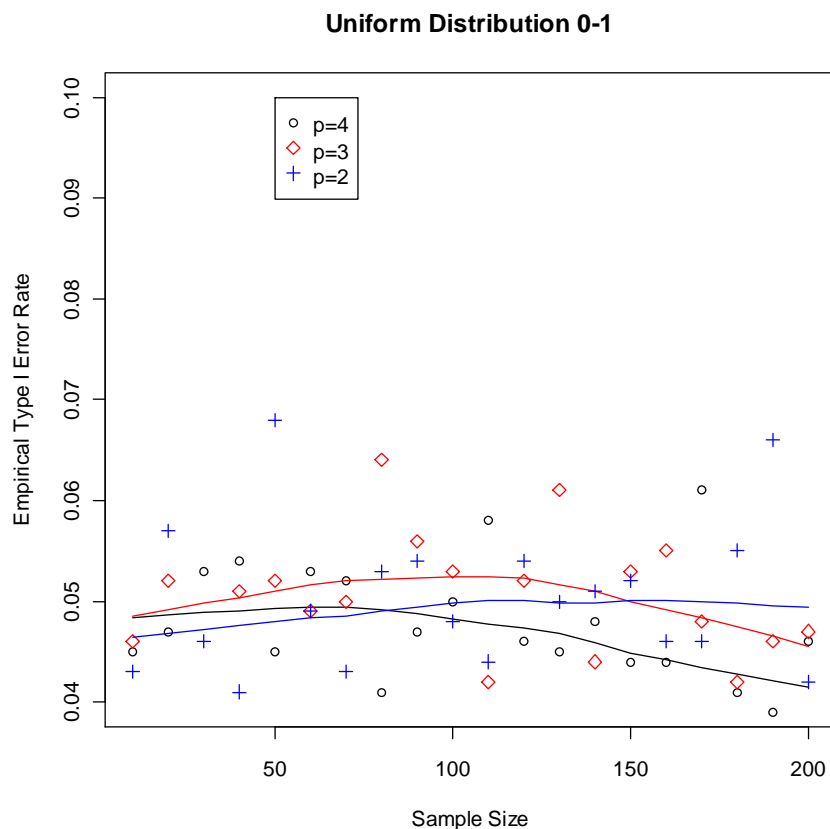
## Multivariate Uniform

In the following simulations data was simulated from a multivariate normal distribution for all three groups. The R software did not have a multivariate uniform distribution simulation function, so in order to simulate this data, univariate uniform distributions were pasted together to form a multivariate data set. In doing this we have to assume that samples within each group are independent and uncorrelated. Essentially, there is zero covariance between each of the dependent variables, but the variance can still be changed to violate some assumptions.

## Simulation V

This first simulation using multivariate uniform data looks at three groups simulated from the same uniform distribution. The uniform distribution allows for data to be uniformly distributed over a certain set of bounds. If the bound are the same, then the mean and variance of the data is the same. In this case each group has a minimum bound equal to zero and a maximum bound equal to one. The only assumption we are violating in this case is the normality assumption because even though the data is multivariate uniform the three groups do have equal covariance structures. Figure 7 is a graph of the empirical type I error rate with sample size on the x-axis.

Figure 8: Uniform 0-1



As seen in figure 8, the empirical type I error rate did not seem to stray very far from our nominal level. The average empirical type I error rates can be found in Table 6 below. Although the error rate for two variables is above 0.05, it is not significantly higher. This difference could just be by random chance. Also, it looks like a couple points in the graph are highly influential in making this average. Overall, the MANOVA procedure seemed to be a little robust against this normality assumption and produced results that stayed very close to our nominal alpha level. If you look at the fitted lines for three and four dependent variables it looks like that there may be a downward trend as sample size increases. After this discovery I re-ran the simulation a couple of times to see if these results could be replicated. After further investigation, sample size did not seem to have an effect on empirical type I error rate for this simulation.

Table 6: Simulation V Results

Number of Variables	Average Empirical Type I Error Rate
4	0.0479
3	0.0498
2	0.0504

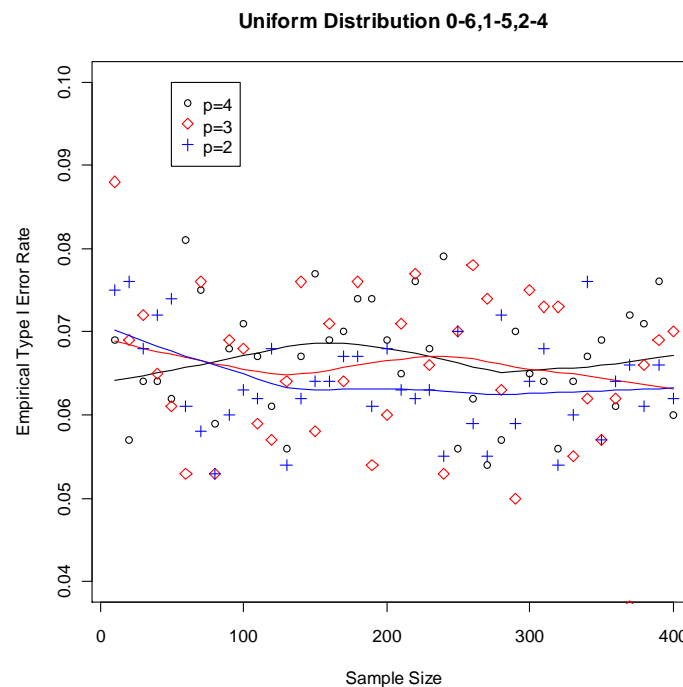
## Simulation VI

Now that we have investigated uniform data with the same bounds for each group, it only seemed natural to see what happened when we changed the bounds between groups. We have to be careful when doing this because we still want the means for each group to be equal, so we must carefully choose the bounds we use. The graph below shows what happened to the type I error rates when we sampled group one with bounds (0, 6), group two with bounds (1, 5), and group three with bounds (2, 4). All of these groups still have the same mean but the variance of the variables in each group changes. Please refer to Table 7 for group information. Due to the way the sampling was done, there is no covariance between variables. Now we are violating two main assumptions: normality and equal covariance structures. Figure 9 explains the effect this has on empirical type I error rate.

Table 7: Variable Information

Group	Variable Mean	Variable Variance
1	3	3
2	3	1.333
3	3	0.333

Figure 9: Uniform, Different Bounds



Here, we see a big stray away from our nominal alpha level, which was expected. We are violating the two main assumptions of MANOVA and would expect the type I error rate to increase. The average rates can be found in Table 8 below. All of these rates are well above

0.05, and we can conclude that when the samples are from a uniform distribution with different bounds for each group, the empirical type I error rate can be above the nominal level.

**Table 8: Simulation VI Results**

Number of Variables	Average Empirical Type I Error Rate
4	0.06665
3	0.06535
2	0.06382

## Functions for Multivariate Methods

Although the R software can easily perform all the multivariate analysis that are necessary for most applications, some of the code is fairly tedious and might be difficult for someone with little to no experience with statistical software. I took that tedious code and condensed it into some very useful functions that would help in-experienced students. The basis for writing these functions was for use in the multivariate statistics class offered at Cal Poly (Stat 419). I wrote four main functions. The first three were fairly fundamental, just some condensing of code into a usable function. These functions included one sample, two sample, and paired Hotellings  $T^2$  test.

One sample Hotelling's  $T^2$  test looks to test for if a population vector is equal to a hypothesized vector. Similarly, a two sample Hotelling's  $T^2$  tests if two population vectors are equal to each other. When you have paired observations for each population, you would then use the paired Hotelling's  $T^2$  test. Please see appendix section C for these functions.

The fourth function that I wrote corresponded to a multivariate method called profile analysis. This function was more complicated and difficult to write than the others. This function can perform profile analysis for k number of groups. Profile analysis a technique used to see how similar different groups are when comparing sequential measurements. There are three main inputs. The first is a data matrix in which the first column is a vector containing the categorical variable that you are comparing by (group, gender, etc.), and the rest of the column vectors contain the sequential measurements in which the means will be tested. The second input pertains to what type of test you want to perform. There are three options: Test="P" for parallelism, Test="C" for coincidence, and Test="F" for flatness. Finally, the third input corresponds to how many groups are being tested. A copy of this function can be found in section D of the appendix.



## Future Projects

My project barely skimmed the surface of what can be done in type I error rate analysis. I looked at particular sets of correlations between variables for each group, but looping over many different sets is easily possible. I also looked at the uniform distribution, but I did not have the time to see how to introduce covariance between the dependent variables. This is another thing that can be looked into in the future. There is also the possibility of looking at other multivariate distributions. We were thinking about introducing a skewed distribution like chi-squared. There is multivariate extension of the chi-squared distribution called the Wishart distribution, but unfortunately I did not have the time look into it. I think that it would be very interesting to see what would happen to the type I error rate when other distributions are investigated.

As well as looking at different violations of assumptions, there are other settings that could be looked at that might affect type I error rate. Although it is not mentioned in my report, I performed some minor analysis corresponding to which type of test statistic was used in MANOVA. All of my simulations were run using the Pillai test statistic, but there are others that are known to help in certain situations. For instance the Wilks test statistic is known to be a little more robust against the normality assumption. These different test statistics can be investigated to see how they affect type I error rate.

## Conclusion

After all of the simulations and analysis had been run, some definite conclusions can be drawn. When the correlations are different across groups, making the covariance matrices differ, the type I error rate can differ from the nominal level. If the correlation is the same between variables for each group, MANOVA assumptions are not violated and the empirical type I error rate stays below our nominal level. However, when each group has different correlation coefficients the error rate was significantly higher. If each group's dependent variables are correlated differently, I would not expect the MANOVA procedure to hold at the nominal alpha level.

When sampling from the same uniform distribution for each group, the type I error rate of MANOVA seemed to hold at our nominal level. Although when sampling different groups from different uniform distributions, the type I error rate was significantly higher than 0.05. This was expected because we are violating two main assumptions of MANOVA: normality and equal covariance structures.

We would have liked to see a sample size effect in the simulations. The sample size effect we would have liked to see was as the sample size increased our type I error rate would decrease. However, we did not see this downward trend we were looking for in the simulations. However, we just looked at the smoothed data curves and did not perform any statistical analysis on a sample size effect. This could be something that could be investigated further in the future.

Another interesting observation was the effect that the number of variables has on type I error rate. In every simulation, except simulation V, the average empirical type I error rates decrease as the number of variables decreases.

In order to further investigate the method of MANOVA, there are many other questions that could be asked. I have just scraped the surface of the possibilities, and I welcome anyone to further investigate to see what they can discover.

## Appendix

### Section A - Simulation Function

```
emp.p.value.normal <- function(data,size1,size2,size3,cor1,cor2,cor3,times)
{
  count=0
  means <- apply(data[,-1],2,mean)
  sigma <- apply(data[,-1],2,sd)
  covar <- cov(data[,-1])
  size <- c(rep(1,times=size1),rep(2,times=size2),rep(3,times=size3))
## Covariance ##
  n=nrow(covar)
  p=ncol(covar)
  covar1=covar
  covar2=covar
  covar3=covar
  for(i in 1:n)
  {
    for(itor in 1:p)
    {
      if(i-itor!=0)
      {
        covar1[i,itor]=cor1*sqrt(covar[i,i]*covar[itor,itor])
        covar2[i,itor]=cor2*sqrt(covar[i,i]*covar[itor,itor])
        covar3[i,itor]=cor3*sqrt(covar[i,i]*covar[itor,itor])
      }
    }
  }

  for( i in 1:times)
  {
    rand.data1 <- rmvnorm(n=size1, mean=means, sigma=covar1)
    rand.data2 <- rmvnorm(n=size2, mean=means, sigma=covar2)
    rand.data3 <- rmvnorm(n=size3, mean=means, sigma=covar3)
    rand.data.final <- cbind(size,rbind(rand.data1,rand.data2,rand.data3))
    rand.manova.obj <- manova(as.matrix(rand.data.final[,-1])~as.factor(rand.data.final[,1]))
    rand.f <- summary(rand.manova.obj)$stats[1, "approx F"]
    if(rand.f >= qf(.05,summary(rand.manova.obj)$stats[1, "num Df"],summary(rand.manova.obj)$stats[1,
"den Df"],lower.tail=F))
    {
      count=count + 1
    }
    ff[i]=rand.f
  }
  count
  emp.p = count/times
  #output= paste("Emperical P-value=", emp.p,"emperical F stat=", ff)
  output=emp.p
  output
}
```

## Section B - Other Simulations

Figure 10: Uniform, Same Bounds (0.30)

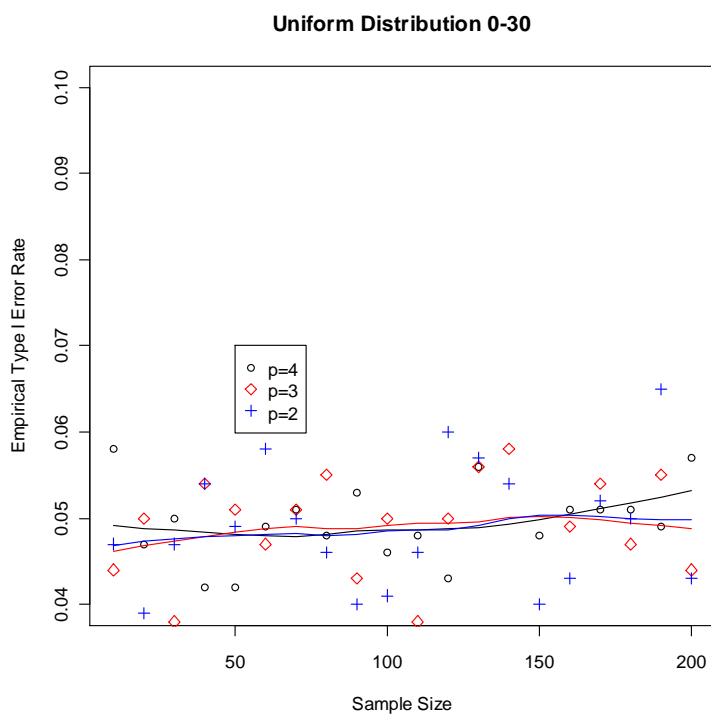


Figure 11: Uniform, Same Bounds (0,10)

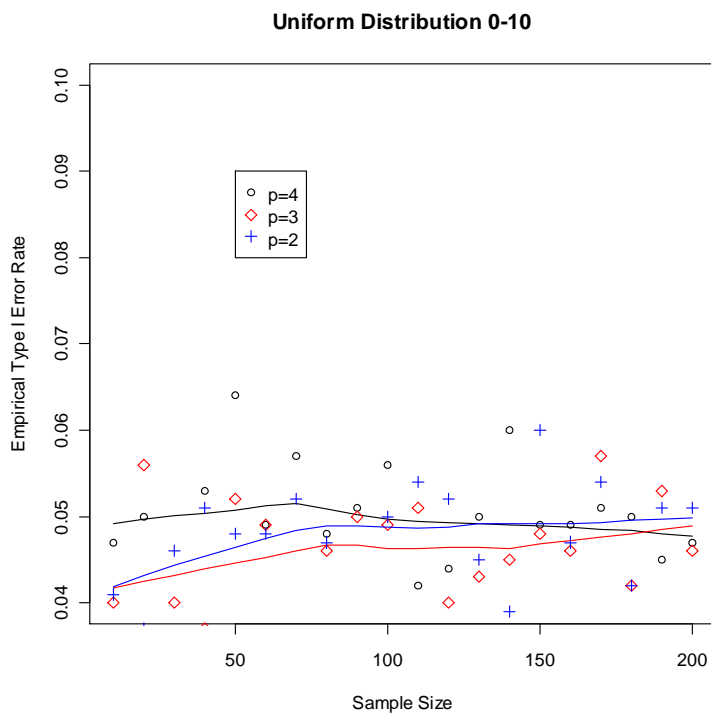
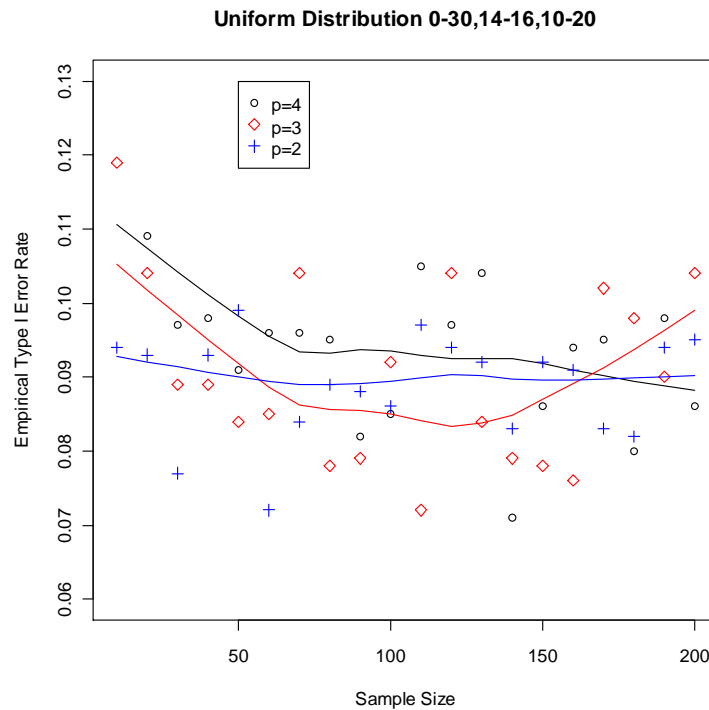


Figure 12: Uniform, Different Bounds (0,30),(14,16),(10,20)



### Section C - Hotellings $T^2$ Functions

```

one.sample.hotelling = function(Data,mu.not)
{
  ### One Sample Hotellings T squared test ###
  ### data: matrix of the 2 vectors you want tested ###
  ### mu.not:vector of length p, with hypothesized values of mu ###

  Data=Data
  n=nrow(Data)
  p=ncol(Data)
  S <- cov(Data)
  y.bar <- apply(Data,2,mean)
  ybar.minus.mu <- y.bar - mu.not
  T.square <- n*t(ybar.minus.mu)%*%solve(S)%*%ybar.minus.mu
  F.stat <- round((n-p)/(p*(n-1))*T.square,digits=3)
  P.value <- round(1-pf(F.stat,p,n-2),digits=5)
  Output <- paste("F=",F.stat,"P-Value=", P.value,sep=" ")
  return(Output)
}

```

```
#####
two.sample.hotelling = function(formula,data)
{
  as.formula(formula)
  as.matrix(data)
  ### Formula: Y~X format, Y is matrix of variables whose means you want tested, using variable
  names for input ###
  ### Data: Name of the data frame ###
  manova.obj <- manova(formula,data=data)
  return(summary(manova.obj))
}
#####

paired.hotelling = function(Data)
{
  ### Data: Data matrix-all y variables first p/2 columns, all x variables [(p/2)-p] ###
  n=nrow(Data)
  col=ncol(Data)
  half=col/2
  D=matrix(0,n,half)
  for(i in 1:half)
  {
    d <- Data[,i]-Data[,half+i]
    D[,i]=d
  }
  paired.Tsq <- manova(D~1)
  return(summary(paired.Tsq,intercept=T,test="Hotelling"))
}
```

### **Section D - K Sample Profile Analysis Function**

```
k.sample.profile = function(Data,Test="P",K=1)
{
  ### Data: first column catagorical variable, next columns, sequential measurements ###
  ### Test: which type of test, p=parallelism,c=coincidence,f=flatness ###
  ### K: number of groups, default is 1 ###
  Data <- as.matrix(Data)
  nrow=nrow(Data)
  ncol=ncol(Data)
  C=matrix(0,ncol-1,ncol)
  for(itor in 1:ncol-1)
  {
    for(itorat in 1:ncol)
    {
      if(itorat-itor==0)
      {
        C[itor,itorat]=1
      }
    }
  }
}
```

```

        if(itorat-itor==1)
        {
            C[itor,itorat]=-1
        }
    }
}
trans.data <- Data%*%t(C)
one.samp.profile <- manova(trans.data~1)
output=(summary(one.samp.profile,intercept=T,test="Hotelling"))

if((Test=="P" | Test=="p") & K!= 1)
{
    crow=ncol-2
    ccol=ncol-1
    C=matrix(0,crow,ccol)
    for(itor in 1:crow)
    {
        for(itorat in 1:ccol)
        {

            if(itorat-itor==0)
            {
                C[itor,itorat]=1
            }
            if(itorat-itor==1)
            {
                C[itor,itorat]=-1
            }
        }
    }
    Y <- as.matrix(Data[,2:ncol])
    trans.data <- Y%*%t(C)
    two.samp.profile <- manova(trans.data~as.factor(Data[,1]))
    output=summary(two.samp.profile,test="Wilks")
}
if((Test=="C" | Test=="c") & K!= 1)
{
    j <- rep(1,times=ncol-1)
    Y <- as.matrix(Data[,2:ncol])
    trans.data <- Y%*%j
    coin.test <- aov(trans.data~as.factor(Data[,1]))
    output=summary(coin.test,test=F)
}

if((Test=="F" | Test=="f") & K!= 1)
{

```

```
crow=ncol-2
ccol=ncol-1
C=matrix(0,crow,ccol)
for(itor in 1:crow)
{
  for(itorat in 1:ccol)
  {
    if(itorat-itor==0)
    {
      C[itor,itorat]=1
    }
    if(itorat-itor==1)
    {
      C[itor,itorat]=-1
    }
  }
}
Y <- as.matrix(Data[,2:ncol])
trans.data <- Y%*%t(C)
test.flat <- manova(trans.data~1+as.factor(Data[,1]))
output=summary(test.flat,intercept=T,test="Hotelling")
}
return(output)
}
```